

EXPRESS MAILING CERTIFICATE

"EXPRESS MAIL" Mailing Label No.: EL084649384 0S

Date of Deposit : August 27, 2001

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, Box Patent Application, Washington, D.C. 20231

Typed or printed name of person signing this certificate:

Signed: Candice Moore



24024

PATENT TRADEMARK OFFICE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
Gorski et al.

) Examiner: Richard Hutson

) Art Unit: 1652

Serial No.: 09/078,465

Filed: May 14, 1998

For: HOMEBOX GENE

) Docket No.: 22311/04015

Assistant Commissioner of Patents

Washington, D.C. 20231

**STATEMENT REGARDING COMPUTER READABLE FORM OF
SEQUENCE LISTING**

Dear Sir:

The computer readable form in the above-described continuation application is identical with the last-filed computer readable form submitted with parent Application No. 09/078,465. In accordance with 37 C.F.R. 1.821(e), please use the last-filed computer readable form of the sequence listing, which was filed in the parent Application on September 23, 1997, as the computer readable form of the sequence listing in the instant application. It is understood that the Patent and Trademark Office will make the necessary change in Application number and filing date for the computer readable form that will be used in the instant application.

094067-08201
T02280-24024

A paper copy of the Sequence Listing is included in the specification of the instant, continuation application.

Respectfully submitted,

Dated: August 27, 2001

By: Pamela A. Docherty
Pamela A. Docherty, Reg. No. 40,591
(216) 622-8416

10/280" E2907650

SUBSTITUTE
SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Gorski, David H.
Walsh, Kenneth
- (ii) TITLE OF INVENTION: Growth Arrest Homeobox Gene
- (iii) NUMBER OF SEQUENCES: 19
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Calfee, Halter, and Griswold
(B) STREET: 800 Superior Avenue
(C) CITY: Cleveland
(D) STATE: Ohio
(E) COUNTRY: U.S.A.
(F) ZIP: 44114-2688
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Golrick, Mary E.
(B) REGISTRATION NUMBER: 34829
(C) REFERENCE/DOCKET NUMBER: 22311/00114
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (216) 622-8200
(B) TELEFAX: (216) 241-0816
(C) TELEX: 980499

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2244 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO

0940573-082701
10/280-8/904660

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 197..1108

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCAAGTGTT TATACGTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT 60
CTGCTCAAAC CCGCGCGGCT TTTACATTAG GAGTGAGTGG GGGAGAGTCC TAGGATTTCT 120
AGTGAAAAGT GACAGCGCTT GGTGGACTTT GGGACCTTCG TGAAGTCTTC TGCTTGAAG 180
CTGAGACTTG CATGCC ATG GAA CAC CCC CTC TTT GGC TGC CTG CGC AGC 229
Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser
1 5 10
CCC CAC GCC ACA GCG CAA GGC TTG CAC CCC TTC TCG CAG TCT TCT CTG 277
Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu
15 20 25
GCC CTC CAT GGA AGA TCT GAC CAC ATG TCC TAC CCC GAA CTC TCC ACA 325
Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr
30 35 40
TCT TCC TCG TCT TGC ATA ATC GCG GGA TAC CCC AAT GAG GAG GGC ATG 373
Ser Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met
45 50 55
TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC CAC CAC CAC CAC 421
Phe Ala Ser Gln His His Arg Gly His His His His His His His His
60 65 70 75
CAT CAC CAC CAC CAG CAG CAG CAG CAC CAG GCT CTG CAA AGC AAC TGG 469
His His His His Gln Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp
80 85 90
CAC CTC CCG CAG ATG TCC TCC CCG CCA AGC GCG GCC CGG CAC AGC CTT 517
His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu
95 100 105
TGC CTG CAG CCT GAT TCC GGA GGG CCC CCG GAG CTG GGG AGC AGC CCT 565
Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro
110 115 120
CCG GTC CTG TGC TCC AAC TCT TCT AGC CTG GGC TCC AGC ACC CCG ACC 613
Pro Val Leu Cys Ser Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr
125 130 135
GGA GCC GCG TGC GCA CCA AGG GAT TAT GGC CGT CAA GCG CTG TCA CCC 661
Gly Ala Ala Cys Ala Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro

00940673 082701

140	145	150	155	
GCA GAA GTG GAG AAG AGA AGT GGC AGC AAA AGA AAA AGC GAC AGT TCA Ala Glu Val Glu Lys Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser 160 165 170				709
GAT TCC CAG GAA GGA AAT TAC AAG TCA GAA GTG AAC AGC AAA CCT AGG Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg 175 180 185				757
AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA Arg Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala 190 195 200				805
GAG TTC GCC CAT CAT AAC TAT CTG ACC AGA CTG AGA AGA TAT GAG ATA Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile 205 210 215				853
GCG GTG AAC CTA GAC CTC ACT GAA AGA CAG GTG AAA GTG TGG TTC CAG Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln 220 225 230 235				901
AAC AGG AGA ATG AAG TGG AAG CGG GTC AAG GGG GGA CAA CAA GGA GCT Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala 240 245 250				949
GCA GCC CGA GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA ACA CTT CTT Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu 255 260 265				997
CCA TCA GAG CTG TCA GGA ATT GGT GCA GCC ACC CTC CAG CAG ACA GGG Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly 270 275 280				1045
GAC TCA CTA GCA AAT GAC GAC AGT CGC GAT AGT GAC CAC AGC TCT GAG Asp Ser Leu Ala Asn Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu 285 290 295				1093
CAC GCA CAC TTA TGATACATAC AGAGACCAGC TCCGTTCTCA GGAAAGCACC His Ala His Leu 300				1145
ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG				1205
CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC				1265
ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA				1325
AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA				1385
AAATTAAATT GCTACCAAGA GCAAACCTCGG TAAGACATTT TGA CTCAAGT TGTCTCCAGA				1445
GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC				1505

05940673-082701
FD2280-E2901650

TGGGCAGGTA TTTGCTTTTG CTTGCACTGA AACTTAAACT GCTATCAAGT TAACCCATGA	1565
AATAGTTTAT CTTGAACAGC CACAGTGCCT GAAATCACCA AGTGGATATA AAATGAACTG	1625
AAATTCTGTA TATATTACTC CTAAGTCATT TTCCTGTCTT CACTAATTTT AGCAAATGCA	1685
TTCATATTAG CTGATGAAAA TAGGCTTTCC CGTGGACAAA TGCAGCCAGC TTCTTGTATT	1745
TTTATACATT TTTTGTGTCAG TCAGAGACAT CAGTATGTGC TTAAGTGTGT TCAAGTAGAG	1805
GAAATGCAGT AGAGTCTGAT AGGACATATT CTTGGTACCA CAGACAAAAC AAATCTTCTG	1865
TTGCATTGAC TATCAACTGC TGCAGATACA TTAGAGAACA CACCTAGCCC CCCTCCAGCC	1925
TCCCTCTGTT ATCGCTCGAA GACATTAGCG TCATAGGCAA GTAGTTACCT TGCCAAATGA	1985
GTCTTGTGTG GCAGATGTCT GATTTTGTAT CTTTAAACTG TTAATGGTAT GTGTCTGCTT	2045
CAGTTAACAG GGAAAAAGAT TTCTTCCTCA TTGTTTATGA TACAAAACCC AAGTGCCAAA	2105
CAAAGCTAGT TCTTCAAGGG ATAGATGAGA AACTGAATGT CTGACAAGTA GACTCAGCGA	2165
AAATACATTA TTTTTCAGAG GCTGTGTATT CATGCAGTAC AAGTCCTTGT ATTTTGTAAG	2225
AAAAAAAGTT AAATAAATG	2244

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 303 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Glu	His	Pro	Leu	Phe	Gly	Cys	Leu	Arg	Ser	Pro	His	Ala	Thr	Ala	
1				5					10					15		
Gln	Gly	Leu	His	Pro	Phe	Ser	Gln	Ser	Ser	Leu	Ala	Leu	His	Gly	Arg	
			20					25					30			
Ser	Asp	His	Met	Ser	Tyr	Pro	Glu	Leu	Ser	Thr	Ser	Ser	Ser	Ser	Cys	
		35					40					45				
Ile	Ile	Ala	Gly	Tyr	Pro	Asn	Glu	Glu	Gly	Met	Phe	Ala	Ser	Gln	His	
	50				55						60					
His	Arg	Gly	His	His	His	His	His	His	His	His	His	His	His	His	Gln	
65				70				75							80	
Gln	Gln	Gln	His	Gln	Ala	Leu	Gln	Ser	Asn	Trp	His	Leu	Pro	Gln	Met	

0940673-082701

85										90					95				
Ser	Ser	Pro	Pro	Ser	Ala	Ala	Arg	His	Ser	Leu	Cys	Leu	Gln	Pro	Asp				
		100						105					110						
Ser	Gly	Gly	Pro	Pro	Glu	Leu	Gly	Ser	Ser	Pro	Pro	Val	Leu	Cys	Ser				
		115					120					125							
Asn	Ser	Ser	Ser	Leu	Gly	Ser	Ser	Thr	Pro	Thr	Gly	Ala	Ala	Cys	Ala				
		130				135					140								
Pro	Arg	Asp	Tyr	Gly	Arg	Gln	Ala	Leu	Ser	Pro	Ala	Glu	Val	Glu	Lys				
145					150				155						160				
Arg	Ser	Gly	Ser	Lys	Arg	Lys	Ser	Asp	Ser	Ser	Asp	Ser	Gln	Glu	Gly				
				165					170					175					
Asn	Tyr	Lys	Ser	Glu	Val	Asn	Ser	Lys	Pro	Arg	Arg	Glu	Arg	Thr	Ala				
			180					185					190						
Phe	Thr	Lys	Glu	Gln	Ile	Arg	Glu	Leu	Glu	Ala	Glu	Phe	Ala	His	His				
		195					200					205							
Asn	Tyr	Leu	Thr	Arg	Leu	Arg	Arg	Tyr	Glu	Ile	Ala	Val	Asn	Leu	Asp				
		210				215					220								
Leu	Thr	Glu	Arg	Gln	Val	Lys	Val	Trp	Phe	Gln	Asn	Arg	Arg	Met	Lys				
225					230					235					240				
Trp	Lys	Arg	Val	Lys	Gly	Gly	Gln	Gln	Gly	Ala	Ala	Ala	Arg	Glu	Lys				
				245					250					255					
Glu	Leu	Val	Asn	Val	Lys	Lys	Gly	Thr	Leu	Leu	Pro	Ser	Glu	Leu	Ser				
			260					265					270						
Gly	Ile	Gly	Ala	Ala	Thr	Leu	Gln	Gln	Thr	Gly	Asp	Ser	Leu	Ala	Asn				
		275					280					285							
Asp	Asp	Ser	Arg	Asp	Ser	Asp	His	Ser	Ser	Glu	His	Ala	His	Leu					
	290					295					300								

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 941 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 33..941

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTCTTCTACC	TGGAACCCGA	AACTTGCATG	CT	ATG	GAA	CAC	CCG	CTC	TTT	GGC		53
				Met	Glu	His	Pro	Leu	Phe	Gly		
				1				5				
TGC	CTG	CGC	AGC	CCT	CAC	GCC	ACG	GCG	CAA	GGC	TTG	101
Cys	Leu	Arg	Ser	Pro	His	Ala	Thr	Ala	Gln	Gly	Leu	
		10					15				20	
CAA	TCC	TCT	CTC	GCC	CTC	CAT	GGA	AGA	TCT	GAC	CAT	149
Gln	Ser	Ser	Leu	Ala	Leu	His	Gly	Arg	Ser	Asp	His	
		25				30				35		
GAG	CTC	TCT	ACT	TCT	TCC	TCA	TCT	TGC	ATA	ATC	GCG	197
Glu	Leu	Ser	Thr	Ser	Ser	Ser	Ser	Cys	Ile	Ile	Ala	
	40				45				50			55
GAA	GAG	GAC	ATG	TTT	GCC	AGC	CAG	CAT	CAC	AGG	GGG	245
Glu	Glu	Asp	Met	Phe	Ala	Ser	Gln	His	His	Arg	Gly	
				60				65				70
CAC	CAC	CAC	CAT	CAC	CAC	CAT	CAG	CAG	CAG	CAG	CAC	293
His	His	His	His	His	His	His	Gln	Gln	Gln	Gln	His	
			75					80				85
ACC	AAC	TGG	CAC	CTC	CCG	CAG	ATG	TCT	TCC	CCA	CCG	341
Thr	Asn	Trp	His	Leu	Pro	Gln	Met	Ser	Ser	Pro	Pro	
		90					95				100	
CAT	AGC	CTC	TGC	CTC	CAG	CCC	GAC	TCT	GGA	GGG	CCC	389
His	Ser	Leu	Cys	Leu	Gln	Pro	Asp	Ser	Gly	Gly	Pro	
	105					110					115	
AGC	AGC	CCG	CCC	GTC	CTG	TGC	TCC	AAC	TCT	TCC	AGC	437
Ser	Ser	Pro	Pro	Val	Leu	Cys	Ser	Asn	Ser	Ser	Leu	
	120				125					130		135
ACC	CCG	ACT	GGG	GCC	GCG	TGC	GCG	CCG	GGG	GAC	TAC	485
Thr	Pro	Thr	Gly	Ala	Ala	Cys	Ala	Pro	Gly	Asp	Tyr	
				140					145			150
CTG	TCA	CCT	GCG	GAG	GCG	GAG	AAG	CGA	AGC	GGC	GGC	533
Leu	Ser	Pro	Ala	Glu	Ala	Glu	Lys	Arg	Ser	Gly	Gly	
			155					160				165

094067-082701

GAC AGC TCA GAC TCC CAG GAA GGA AAT TAC AAG TCA GAA GTC AAC AGC 581
 Asp Ser Ser Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser
 170 175 180

AAA CCC AGG AAA GAA AGG ACA GCA TTT ACC AAA GAG CAA ATC AGA GAA 629
 Lys Pro Arg Lys Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu
 185 190 195

CTT GAA GCA GAA TTT GCC CAT CAT AAT TAT CTC ACC AGA CTG AGG CGA 677
 Leu Glu Ala Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg
 200 205 210 215

TAC GAG ATA GCA GTG AAT CTG GAT CTC ACT GAA AGA CAG GTA AAA GTC 725
 Tyr Glu Ile Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val
 220 225 230

TGG TTC CAA AAC AGG CGG ATG AAG TGG AAG AGG GTA AAG GGT GGA CAG 773
 Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln
 235 240 245

CAA GGA GCT GCG GCT CGG GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA 821
 Gln Gly Ala Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly
 250 255 260

ACA CTT CTC CCA TCA GAG CTG TCG GGA ATT GGT GCA GCC ACC CTC CAG 869
 Thr Leu Leu Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln
 265 270 275

CAA ACA GGG GAC TCT ATA GCA AAT GAA GAC AGT CAC GAC AGT GAC CAC 917
 Gln Thr Gly Asp Ser Ile Ala Asn Glu Asp Ser His Asp Ser Asp His
 280 285 290 295

AGC TCA GAG CAC GCC CAC CTC TGA 941
 Ser Ser Glu His Ala His Leu
 300

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 302 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser Pro His Ala Thr Ala
 1 5 10 15

Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu Ala Leu His Gly Arg
 20 25 30

00940673 002701

0940673-082704

Ser	Asp	His	Met	Ser	Tyr	Pro	Glu	Leu	Ser	Thr	Ser	Ser	Ser	Ser	Cys
		35					40					45			
Ile	Ile	Ala	Gly	Tyr	Pro	Asn	Glu	Glu	Asp	Met	Phe	Ala	Ser	Gln	His
	50					55					60				
His	Arg	Gly	His	His	His	His	His	His	His	His	His	His	His	Gln	Gln
65					70					75					80
Gln	Gln	His	Gln	Ala	Leu	Gln	Thr	Asn	Trp	His	Leu	Pro	Gln	Met	Ser
				85					90					95	
Ser	Pro	Pro	Ser	Ala	Ala	Arg	His	Ser	Leu	Cys	Leu	Gln	Pro	Asp	Ser
			100					105					110		
Gly	Gly	Pro	Pro	Glu	Leu	Gly	Ser	Ser	Pro	Pro	Val	Leu	Cys	Ser	Asn
		115					120					125			
Ser	Ser	Ser	Leu	Gly	Ser	Ser	Thr	Pro	Thr	Gly	Ala	Ala	Cys	Ala	Pro
	130					135					140				
Gly	Asp	Tyr	Gly	Arg	Gln	Ala	Leu	Ser	Pro	Ala	Glu	Ala	Glu	Lys	Arg
145					150					155					160
Ser	Gly	Gly	Lys	Arg	Lys	Ser	Asp	Ser	Ser	Asp	Ser	Gln	Glu	Gly	Asn
				165					170					175	
Tyr	Lys	Ser	Glu	Val	Asn	Ser	Lys	Pro	Arg	Lys	Glu	Arg	Thr	Ala	Phe
			180					185					190		
Thr	Lys	Glu	Gln	Ile	Arg	Glu	Leu	Glu	Ala	Glu	Phe	Ala	His	His	Asn
		195					200					205			
Tyr	Leu	Thr	Arg	Leu	Arg	Arg	Tyr	Glu	Ile	Ala	Val	Asn	Leu	Asp	Leu
	210					215					220				
Thr	Glu	Arg	Gln	Val	Lys	Val	Trp	Phe	Gln	Asn	Arg	Arg	Met	Lys	Trp
225					230					235					240
Lys	Arg	Val	Lys	Gly	Gly	Gln	Gln	Gly	Ala	Ala	Ala	Arg	Glu	Lys	Glu
				245					250					255	
Leu	Val	Asn	Val	Lys	Lys	Gly	Thr	Leu	Leu	Pro	Ser	Glu	Leu	Ser	Gly
			260					265					270		
Ile	Gly	Ala	Ala	Thr	Leu	Gln	Gln	Thr	Gly	Asp	Ser	Ile	Ala	Asn	Glu
		275					280					285			
Asp	Ser	His	Asp	Ser	Asp	His	Ser	Ser	Glu	His	Ala	His	Leu		
	290					295					300				

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 6
- (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 21
- (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 24
- (D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AARATWTGGT TYCARAAYMG WMGWATGAA

29

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 4
- (D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

09940673-082704
FDZ280 E2904560

TCAWARRTGW GCRTGYTC

18

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCGCGCAGAT CTCACTGAAA GACAGGTAAA

30

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TTTACCTGTC TTTCAGTGAG

20

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

094067.00704
10/28/90 12:50

[Faint, illegible text from bleed-through]

[illegible]

[Faint, illegible text from bleed-through]

[Faint, illegible text from bleed-through]

[Faint, illegible text from bleed-through]

[illegible]

[Faint, illegible text from bleed-through]

[illegible][illegible]

[Faint, illegible text from bleed-through]

[illegible][illegible]

[Faint, illegible text from bleed-through]

[illegible][illegible][illegible][illegible][illegible]

[Faint, illegible text from bleed-through]

[illegible]

[Faint, illegible text from bleed-through]

[Faint, illegible text from bleed-through]

THE **NEW** **YORK** **PUBLIC** **LIBRARY**

[illegible][illegible]

THE **NEW** **YORK** **PUBLIC** **LIBRARY**

[illegible]

THE **NEW** **YORK** **PUBLIC** **LIBRARY**

THE **NEW** **YORK** **PUBLIC** **LIBRARY**

THE **NEW** **YORK** **PUBLIC** **LIBRARY**

[illegible][illegible][illegible]

THE **NEW** **YORK** **PUBLIC** **LIBRARY**

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CTGGTTCGGC CCACCTCTGA AGGTTCCAGA ATCGATAG

38

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGAGACTTCC AAGGTCTTAG CTATCACTTA AGCAC

35

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCGCGCGTCG ACGAACACCC CCTCTTTGGC

30

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid

0540573.082701
10/28/95 15:50

1890
 1891
 1892
 1893
 1894
 1895
 1896
 1897
 1898
 1899
 1900
 1901
 1902
 1903
 1904
 1905
 1906
 1907
 1908
 1909
 1910
 1911
 1912
 1913
 1914
 1915
 1916
 1917
 1918
 1919
 1920
 1921
 1922
 1923
 1924
 1925
 1926
 1927
 1928
 1929
 1930
 1931
 1932
 1933
 1934
 1935
 1936
 1937
 1938
 1939
 1940
 1941
 1942
 1943
 1944
 1945
 1946
 1947
 1948
 1949
 1950
 1951
 1952
 1953
 1954
 1955
 1956
 1957
 1958
 1959
 1960
 1961
 1962
 1963
 1964
 1965
 1966
 1967
 1968
 1969
 1970
 1971
 1972
 1973
 1974
 1975
 1976
 1977
 1978
 1979
 1980
 1981
 1982
 1983
 1984
 1985
 1986
 1987
 1988
 1989
 1990
 1991
 1992
 1993
 1994
 1995
 1996
 1997
 1998
 1999
 2000
 2001
 2002
 2003
 2004
 2005
 2006
 2007
 2008
 2009
 2010
 2011
 2012
 2013
 2014
 2015
 2016
 2017
 2018
 2019
 2020
 2021
 2022
 2023
 2024
 2025
 2026
 2027
 2028
 2029
 2030
 2031
 2032
 2033
 2034
 2035
 2036
 2037
 2038
 2039
 2040
 2041
 2042
 2043
 2044
 2045
 2046
 2047
 2048
 2049
 2050
 2051
 2052
 2053
 2054
 2055
 2056
 2057
 2058
 2059
 2060
 2061
 2062
 2063
 2064
 2065
 2066
 2067
 2068
 2069
 2070
 2071
 2072
 2073
 2074
 2075
 2076
 2077
 2078
 2079
 2080
 2081
 2082
 2083
 2084
 2085
 2086
 2087
 2088
 2089
 2090
 2091
 2092
 2093
 2094
 2095
 2096
 2097
 2098
 2099
 2100
 2101
 2102
 2103
 2104
 2105
 2106
 2107
 2108
 2109
 2110
 2111
 2112
 2113
 2114
 2115
 2116
 2117
 2118
 2119
 2120
 2121
 2122
 2123
 2124
 2125
 2126
 2127
 2128
 2129
 2130
 2131
 2132
 2133
 2134
 2135
 2136
 2137
 2138
 2139
 2140
 2141
 2142
 2143
 2144
 2145
 2146
 2147
 2148
 2149
 2150
 2151
 2152
 2153
 2154
 2155
 2156
 2157
 2158
 2159
 2160
 2161
 2162
 2163
 2164
 2165
 2166
 2167
 2168
 2169
 2170
 2171
 2172
 2173
 2174
 2175
 2176
 2177
 2178
 2179
 2180
 2181
 2182
 2183
 2184
 2185
 2186
 2187
 2188
 2189
 2190
 2191
 2192
 2193
 2194
 2195
 2196
 2197
 2198
 2199
 2200
 2201
 2202
 2203
 2204
 2205
 2206
 2207
 2208
 2209
 2210
 2211
 2212
 2213
 2214
 2215
 2216
 2217
 2218
 2219
 2220
 2221
 2222
 2223
 2224
 2225
 2226
 2227
 2228
 2229
 2230
 2231
 2232
 2233
 2234
 2235
 2236
 2237
 2238
 2239
 2240
 2241
 2242
 2243
 2244
 2245
 2246
 2247
 2248
 2249
 2250
 2251
 2252
 2253
 2254
 2255
 2256
 2257
 2258
 2259
 2260
 2261
 2262
 2263
 2264
 2265
 2266
 2267
 2268
 2269
 2270
 2271
 2272
 2273
 2274
 2275
 2276
 2277
 2278
 2279
 2280
 2281
 2282
 2283
 2284
 2285
 2286
 2287
 2288
 2289
 2290
 2291
 2292
 2293
 2294
 2295
 2296
 2297
 2298
 2299
 2300
 2301
 2302
 2303
 2304
 2305
 2306
 2307
 2308
 2309
 2310
 2311
 2312
 2313
 2314
 2315
 2316
 2317
 2318
 2319
 2320
 2321
 2322
 2323
 2324
 2325
 2326
 2327
 2328
 2329
 2330
 2331
 2332
 2333
 2334
 2335
 2336
 2337
 2338
 2339
 2340
 2341
 2342
 2343
 2344

(iv) ANTI-SENSE: NO

30

(iv) ANTI-SENSE: NO

25

(iv) ANTI-SENSE: NO

25

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGATGGCATG GACTGTGGTC ATGA

24

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGATGGCATG GACTGTGGTC ATGA

24

094057.0876
"2280" 52904560

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Gorski, David H.
Walsh, Kenneth

(ii) TITLE OF INVENTION: Growth Arrest Homeobox Gene

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Calfee, Halter, and Griswold
(B) STREET: 800 Superior Avenue
(C) CITY: Cleveland
(D) STATE: Ohio
(E) COUNTRY: U.S.A.
(F) ZIP: 44114-2688

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Golrick, Mary E.
(B) REGISTRATION NUMBER: 34829
(C) REFERENCE/DOCKET NUMBER: 22311/00114

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (216) 622-8200
(B) TELEFAX: (216) 241-0816
(C) TELEX: 980499

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2244 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 197..1108

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCAAGTGTT TATACGTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT
60

22311/00114

34

CTGCTCAAAC CCGCGCGGCT TTTACATTAG GAGTGAGTGG GGGAGAGTCC TAGGATTTCT
120

AGTGAAAAGT GACAGCGCTT GGTGGACTTT GGGACCTTCG TGAAGTCTTC TGCTTGGAAG
180

CTGAGACTTG CATGCC ATG GAA CAC CCC CTC TTT GGC TGC CTG CGC AGC
229

Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser
1 5 10

CCC CAC GCC ACA GCG CAA GGC TTG CAC CCC TTC TCG CAG TCT TCT CTG
277

Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu
15 15 20 25

GCC CTC CAT GGA AGA TCT GAC CAC ATG TCC TAC CCC GAA CTC TCC ACA
325

Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr
30 35 40

TCT TCC TCG TCT TGC ATA ATC GCG GGA TAC CCC AAT GAG GAG GGC ATG
373

Ser Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met
45 50 55

TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC CAC CAC CAC
421

Phe Ala Ser Gln His His Arg Gly His His His His His His His
60 65 70 75

CAT CAC CAC CAC CAG CAG CAG CAG CAC CAG GCT CTG CAA AGC AAC TGG
469

His His His His Gln Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp
80 85 90

CAC CTC CCG CAG ATG TCC TCC CCG CCA AGC GCG GCC CGG CAC AGC CTT
517

His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu
95 100 105

TGC CTG CAG CCT GAT TCC GGA GGG CCC CCG GAG CTG GGG AGC AGC CCT
565

Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro
110 115 120

CCG GTC CTG TGC TCC AAC TCT TCT AGC CTG GGC TCC AGC ACC CCG ACC
613

Pro Val Leu Cys Ser Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr
125 130 135

GGA GCC GCG TGC GCA CCA AGG GAT TAT GGC CGT CAA GCG CTG TCA CCC
661

Gly Ala Ala Cys Ala Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro
140 145 150 155

GCA GAA GTG GAG AAG AGA AGT GGC AGC AAA AGA AAA AGC GAC AGT TCA
709

Ala Glu Val Glu Lys Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser
160 165 170

GAT TCC CAG GAA GGA AAT TAC AAG TCA GAA GTG AAC AGC AAA CCT AGG
757

Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg
175 180 185

AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA

10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540 550 560 570 580 590 600 610 620 630 640 650 660 670 680 690 700 710 720 730 740 750 760 770 780 790 800 810 820 830 840 850 860 870 880 890 900 910 920 930 940 950 960 970 980 990 1000

35

Arg Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala
190 195 200

5 GAG TTC GCC CAT CAT AAC TAT CTG ACC AGA CTG AGA AGA TAT GAG ATA
853
Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile
205 210 215

10 GCG GTG AAC CTA GAC CTC ACT GAA AGA CAG GTG AAA GTG TGG TTC CAG
901
Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln
220 225 230 235

15 AAC AGG AGA ATG AAG TGG AAG CGG GTC AAG GGG GGA CAA CAA GGA GCT
949
Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala
240 245 250

20 GCA GCC CGA GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA ACA CTT CTT
997
Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu
255 260 265

CCA TCA GAG CTG TCA GGA ATT GGT GCA GCC ACC CTC CAG CAG ACA GGG
1045
Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly
270 275 280

GAC TCA CTA GCA AAT GAC GAC AGT CGC GAT AGT GAC CAC AGC TCT GAG
1093
Asp Ser Leu Ala Asn Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu
285 290 295

CAC GCA CAC TTA TGATACATAC AGAGACCAGC TCCGTTCTCA GGAAAGCACC
1145
His Ala His Leu
300

ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG
1205

CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC
1265

ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA
1325

AAATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA
1385

AAATTAAATT GCTACCAAGA GCAAACCTCGG TAAGACATTT TGA CTCAAGT TGTCTCCAGA
1445

GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC
1505

TGGGCAGGTA TTTGCTTTTG CTTGCACTGA AACTTAACT GCTATCAAGT TAACCCATGA
1565

AATAGTTTAT CTTGAACAGC CACAGTGCCT GAAATCACCA AGTGGATATA AAATGAACTG
1625

AAATTCTGTA TATATTACTC CTAAGTCATT TTCCTGTCTT CACTAATTTT AGCAAATGCA
1685

TTCATATTAG CTGATGAAAA TAGGCTTTCC CGTGGACAAA TGCAGCCAGC TTCTTGTATT

1745

TTTATACATT TTTTGTGTCAG TCAGAGACAT CAGTATGTGC TTACTTGTGT TCAAGTAGAG
1805

GAAATGCAGT AGAGTCTGAT AGGACATATT CTTGGTACCA CAGACAAAAC AAATCTTCTG
1865

TTGCATTGAC TATCAACTGC TGCAGATACA TTAGAGAACA CACCTAGCCC CCCTCCAGCC
1925

TCCCTCTGTT ATCGCTCGAA GACATTAGCG TCATAGGCAA GTAGTTACCT TGCCAAATGA
1985

GTCTTGTGTG GCAGATGTCT GATTTTGTAT CTTTAACTG TTAATGGTAT GTGTCTGCTT
2045

CAGTTAACAG GGAAAAAGAT TTCTTCCTCA TTGTTTATGA TACAAAACCC AAGTGCCAAA
2105

CAAAGCTAGT TCTTCAAGGG ATAGATGAGA AACTGAATGT CTGACAAGTA GACTCAGCGA
2165

AAATACATTA TTTTTCAGAG GCTGTGTATT CATGCAGTAC AAGTCCTTGT ATTTTGTAAG
2225

AAAAAAAAGT AAATAAATG
2244

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 303 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Glu	His	Pro	Leu	Phe	Gly	Cys	Leu	Arg	Ser	Pro	His	Ala	Thr	Ala	1	5	10	15
Gln	Gly	Leu	His	Pro	Phe	Ser	Gln	Ser	Ser	Leu	Ala	Leu	His	Gly	Arg	20	25	30	
Ser	Asp	His	Met	Ser	Tyr	Pro	Glu	Leu	Ser	Thr	Ser	Ser	Ser	Ser	Cys	35	40	45	
Ile	Ile	Ala	Gly	Tyr	Pro	Asn	Glu	Glu	Gly	Met	Phe	Ala	Ser	Gln	His	50	55	60	
His	Arg	Gly	His	His	His	His	His	His	His	His	His	His	His	His	Gln	65	70	75	80
Gln	Gln	Gln	His	Gln	Ala	Leu	Gln	Ser	Asn	Trp	His	Leu	Pro	Gln	Met	85	90	95	
Ser	Ser	Pro	Pro	Ser	Ala	Ala	Arg	His	Ser	Leu	Cys	Leu	Gln	Pro	Asp	100	105	110	
Ser	Gly	Gly	Pro	Pro	Glu	Leu	Gly	Ser	Ser	Pro	Pro	Val	Leu	Cys	Ser	115	120	125	
Asn	Ser	Ser	Ser	Leu	Gly	Ser	Ser	Thr	Pro	Thr	Gly	Ala	Ala	Cys	Ala	130	135	140	

22311/00114

37

Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro Ala Glu Val Glu Lys
145 150 155 160

Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser Asp Ser Gln Glu Gly
165 170 175

Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg Arg Glu Arg Thr Ala
180 185 190

Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His
195 200 205

Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile Ala Val Asn Leu Asp
210 215 220

Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln Asn Arg Arg Met Lys
225 230 235 240

Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala Ala Ala Arg Glu Lys
245 250 255

Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu Pro Ser Glu Leu Ser
260 265 270

Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly Asp Ser Leu Ala Asn
275 280 285

Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu His Ala His Leu
290 295 300

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 941 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 33..941

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTCTTCTACC TGGAACCCGA AACTTGCATG CT ATG GAA CAC CCG CTC TTT GGC
53

Met Glu His Pro Leu Phe Gly
1 5

TGC CTG CGC AGC CCT CAC GCC ACG GCG CAA GGC TTG CAC CCG TTC TCC
101

Cys Leu Arg Ser Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser
10 15 20

CAA TCC TCT CTC GCC CTC CAT GGA AGA TCT GAC CAT ATG TCT TAC CCC
149

Gln Ser Ser Leu Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro
25 30 35

GAG CTC TCT ACT TCT TCC TCA TCT TGC ATA ATC GCG GGA TAC CCC AAC
 197
 Glu Leu Ser Thr Ser Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn
 40 45 50 55
 GAA GAG GAC ATG TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC
 245
 Glu Glu Asp Met Phe Ala Ser Gln His His Arg Gly His His His His
 60 65 70
 CAC CAC CAC CAT CAC CAC CAT CAG CAG CAG CAG CAC CAG GCT CTG CAA
 293
 His His His His His His His Gln Gln Gln Gln His Gln Ala Leu Gln
 75 80 85
 ACC AAC TGG CAC CTC CCG CAG ATG TCT TCC CCA CCG AGT GCG GCT CGG
 341
 Thr Asn Trp His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg
 90 95 100
 CAT AGC CTC TGC CTC CAG CCC GAC TCT GGA GGG CCC CCA GAG TTG GGG
 389
 His Ser Leu Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly
 105 110 115
 AGC AGC CCG CCC GTC CTG TGC TCC AAC TCT TCC AGC TTG GGC TCC AGC
 437
 Ser Ser Pro Pro Val Leu Cys Ser Asn Ser Ser Ser Leu Gly Ser Ser
 120 125 130 135
 ACC CCG ACT GGG GCC GCG TGC GCG CCG GGG GAC TAC GGC CGC CAG GCA
 485
 Thr Pro Thr Gly Ala Ala Cys Ala Pro Gly Asp Tyr Gly Arg Gln Ala
 140 145 150
 CTG TCA CCT GCG GAG GCG GAG AAG CGA AGC GGC GGC AAG AGG AAA AGC
 533
 Leu Ser Pro Ala Glu Ala Glu Lys Arg Ser Gly Gly Lys Arg Lys Ser
 155 160 165
 GAC AGC TCA GAC TCC CAG GAA GGA AAT TAC AAG TCA GAA GTC AAC AGC
 581
 Asp Ser Ser Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser
 170 175 180
 AAA CCC AGG AAA GAA AGG ACA GCA TTT ACC AAA GAG CAA ATC AGA GAA
 629
 Lys Pro Arg Lys Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu
 185 190 195
 CTT GAA GCA GAA TTT GCC CAT CAT AAT TAT CTC ACC AGA CTG AGG CGA
 677
 Leu Glu Ala Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg
 200 205 210 215
 TAC GAG ATA GCA GTG AAT CTG GAT CTC ACT GAA AGA CAG GTA AAA GTC
 725
 Tyr Glu Ile Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val
 220 225 230
 TGG TTC CAA AAC AGG CGG ATG AAG TGG AAG AGG GTA AAG GGT GGA CAG
 773
 Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln
 235 240 245
 CAA GGA GCT GCG GCT CGG GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA
 821

Gln Gly Ala Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly
 250 255 260

ACA CTT CTC CCA TCA GAG CTG TCG GGA ATT GGT GCA GCC ACC CTC CAG
 869

Thr Leu Leu Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln
 265 270 275

CAA ACA GGG GAC TCT ATA GCA AAT GAA GAC AGT CAC GAC AGT GAC CAC
 917

Gln Thr Gly Asp Ser Ile Ala Asn Glu Asp Ser His Asp Ser Asp His
 280 285 290 295

AGC TCA GAG CAC GCC CAC CTC TGA
 941

Ser Ser Glu His Ala His Leu
 300

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 302 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser Pro His Ala Thr Ala
 1 5 10 15

Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu Ala Leu His Gly Arg
 20 25 30

Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr Ser Ser Ser Cys
 35 40 45

Ile Ile Ala Gly Tyr Pro Asn Glu Glu Asp Met Phe Ala Ser Gln His
 50 55 60

His Arg Gly His His His His His His His His His His Gln Gln
 65 70 75 80

Gln Gln His Gln Ala Leu Gln Thr Asn Trp His Leu Pro Gln Met Ser
 85 90 95

Ser Pro Pro Ser Ala Ala Arg His Ser Leu Cys Leu Gln Pro Asp Ser
 100 105 110

Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro Pro Val Leu Cys Ser Asn
 115 120 125

Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr Gly Ala Ala Cys Ala Pro
 130 135 140

Gly Asp Tyr Gly Arg Gln Ala Leu Ser Pro Ala Glu Ala Glu Lys Arg
 145 150 155 160

Ser Gly Gly Lys Arg Lys Ser Asp Ser Ser Asp Ser Gln Glu Gly Asn
 165 170 175

Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg Lys Glu Arg Thr Ala Phe
 180 185 190

Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His Asn

22311/00114

40

195

200

205

Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile Ala Val Asn Leu Asp Leu
210 215 220

Thr Glu Arg Gln Val Lys Val Trp Phe Gln Asn Arg Arg Met Lys Trp
225 230 235 240

Lys Arg Val Lys Gly Gly Gln Gln Gly Ala Ala Ala Arg Glu Lys Glu
245 250 255

Leu Val Asn Val Lys Lys Gly Thr Leu Leu Pro Ser Glu Leu Ser Gly
260 265 270

y
Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly Asp Ser Ile Ala Asn Glu
275 280 285

Asp Ser His Asp Ser Asp His Ser Ser Glu His Ala His Leu
290 295 300

FD-220" 2-904580